

Optimality of Radio Power Control via Fast-Lipschitz Optimization

Martin Jakobsson and Carlo Fischione

Abstract

Fixed point algorithms play an important role to compute feasible solutions to the radio power control problems in wireless networks. Although these algorithms are shown to converge to the fixed points that give feasible problem solutions, the solutions often lack notion of problem optimality. This paper reconsiders well known fixed point algorithms such as those with standard and type-II standard interference functions, and investigates the conditions under which they give optimal power control solutions by the recently proposed Fast-Lipschitz optimization framework. When the qualifying conditions of Fast-Lipschitz optimization apply, it is established that the fixed points are the optimal solutions of radio power optimization problems. The analysis is performed by a logarithmic transformation of variables that gives problems treatable within the Fast-Lipschitz framework. It is shown how the logarithmic problem constraints are contractive by the standard or type-II standard assumptions on the original power control problem, and how a set of cost functions fulfill the Fast-Lipschitz qualifying conditions. The analysis on non monotonic interference function allows to establish a new qualifying condition for Fast-Lipschitz optimization. The results are illustrated by considering power control problems with standard interference function, problems with type-II standard interference functions, and a case of sub-homogeneous power control problems. It is concluded that Fast-Lipschitz optimization may play an important role in many resource allocation problems in wireless networks.

I. INTRODUCTION

Radio power control is one of the essential radio resource management techniques in wireless networks. The power control problem faces a tradeoff between saving power and having high enough level of power. It is important to control the transmit radio powers to

The authors are with the Department of Automatic Control, KTH Royal Institute of Technology, Sweden. Email: {mjakobss,carlofi}@kth.se.

The work of the authors is sponsored by the EU projects NoE HYCON2 and Hydrobionets.

avoid interferences to undesired receivers and save the energy of the transmitters. Meanwhile, it is important also to use adequate levels of power to make sure that the transmitted signals can overcome the attenuation of the wireless channel and the interference caused by other transmitters.

Power control in wireless communication is a particularly successful instance of distributed optimization over networks. Specifically, in wireless networks a link is associated to one pair of nodes where a node is a transmitter and the other is a receiver. Suppose there are n transmitter-receiver pairs. Let p_i be the radio power of transmit node i , for $i = 1, \dots, n$. Note that the index i is used both for a transmitter and a receiver, so *transmitter i* and *receiver i* are two different nodes that are paired to communicate. One of the simplest examples of radio power control considers the gain matrix \mathbf{G} , where G_{ij} is the channel attenuation from the transmit node j to the receiver node i . In addition to the useful signal $G_{ii}p_i$ from transmitter i , receiver i will also receive a background noise η_i plus the interference of all other transmitters, $\sum_{j \neq i} G_{ij}p_j$. For the communication attempt of transmit node i to be successful, the signal to (interference and) noise ratio (SNR) at receiver node i must be higher than some threshold τ_i ,

$$\frac{G_{ii}p_i}{\sum_{j \neq i} G_{ij}p_j + \eta_i} \geq \tau_i. \quad (1)$$

If the transmit power of all links are collected in the vector $\mathbf{p} = [p_1, p_2, \dots, p_n]^T$, the requirement (1) can be rewritten as

$$p_i \geq \mathcal{I}_i(\mathbf{p}) \triangleq \frac{\tau_i}{G_{ii}} \left(\sum_{j \neq i} G_{ij}p_j + \eta_i \right). \quad (2)$$

Using the vocabulary of [1], we will refer to $\mathcal{I}_i(\mathbf{p})$ as the *interference function* of transmitter i . This affine version of $\mathcal{I}_i(\mathbf{p})$ is the simplest and best studied type of interference function, and it is often the basis for extensions or modifications by other types of interference functions. The focus on achieving some minimum SNR in (1) is justified because many other measures of the quality of service are increasing functions of the SNR [2].

There are a number of ways of using the interference function in setting power control optimization problems:

- maximization of the SNR (i.e., quality of service) of the network, subject to power constraints;
- minimization of the power consumption subject to SNR constraints;
- maximization of some network utility function (e.g., throughput) of the network, subject to power constraints.

Early works on distributed power control in wireless networks have followed the first approach, and try to maximize the smallest SNR of the network [3, 4]. With the inclusion of receiver noise in [5], the focus has shifted to the second approach, with the goal of minimizing the radio powers p_i while maintaining a minimum SNR τ_i at each receiver, i.e.,

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{p} \\ \text{s.t.} \quad & p_i \geq \mathcal{I}_i(\mathbf{p}) \quad \forall i. \end{aligned} \quad (3)$$

This line of work has later been generalized to the framework of *standard* interference functions by Yates [1]. When problem (3) above is feasible, and the functions $\mathcal{I}_i(\mathbf{p})$ are standard (see Definition 9), the unique optimal solution to (3) is given by the fixed point of the iteration

$$p_i^{k+1} := \mathcal{I}_i(\mathbf{p}^k), \quad (4)$$

or, in vector form, $\mathbf{p}^{k+1} := \mathcal{I}(\mathbf{p}^k)$ where

$$\mathcal{I}(\mathbf{p}^k) \triangleq \begin{bmatrix} \mathcal{I}_1(\mathbf{p}^k) & \dots & \mathcal{I}_n(\mathbf{p}^k) \end{bmatrix}^T.$$

The computation of the optimal solution for problem (3) by these iterations is much simpler than using the classical parallelization and decomposition methods of distributed optimization [6]. This is because there is no longer a need to centrally collect, compute and redistribute the primal and dual variables of the problem due to that $\mathcal{I}_i(\mathbf{p}^k)$ can be known or estimated locally at receiver i [1, 5]. Even in a centralized setting, iteration (4) is simpler than traditional distributed optimization methods, since no dual variables need to be stored and manipulated. The iterations require only that every receiver node successively updates using local knowledge of the function (interference function) of other nodes' current decision variables (radio powers). Another advantage is that convergence is obtained even though such a knowledge is delayed, i.e., the decision variables p_j^k of other nodes are updated with some delay [6, 7].

Given the advantages mentioned above, there is a number of studies in the literature where radio power control algorithms have been proposed by considering iterations similar to (4) [8–13]. In these approaches, the interference functions are not necessarily standard, and the focus has been in studying the convergence of iteration (4) to a fixed point rather than the meaning of the fixed point in terms of optimality for problem (3). One of the reasons is that optimal power control problems of the form (3) with or without standard interference functions may be non convex and non linear, which makes it very hard the characterization of optimality. Therefore, the natural question that we would like to answer in the paper

is whether there are conditions ensuring the optimality of existing fixed point radio power control algorithms.

Fast-Lipschitz optimization is a recently proposed framework that is motivated by such a question [14?]. In particular, Fast-Lipschitz optimization is a natural generalization of the interference function approach on how to solve distributed optimization problems over wireless networks by using fixed point iterations similar to (4). Fast-Lipschitz does not need constraints that are standard, and can have a more general objective functions than the one in problem (3). The main characteristic of a Fast-Lipschitz problem is that the optimal point is given by the fixed point of the constraints, a result that is in general very difficult to establish. In this paper we investigate under which conditions a general power control problem falls under the Fast-Lipschitz framework, whereby the fixed points of iteration (4) are optimal.

The remainder of this paper is organized as follows: In Subsection I.A we discuss the related work and in Subsection I.B notation is introduced. The problem formulation is given in Section II, and for the sake of self-containment we give a brief definition of Fast-Lipschitz optimization in Section III. In Section IV.A we present preliminary results on two-sided scalable functions, which we use to examine the relation between standard and type-II standard functions in sections IV.B and IV.C respectively. Section V gives an example of Fast-Lipschitz optimization applied to a problem that is neither standard, nor type-II standard. Finally, the paper is concluded in Section VI.

A. *Related work*

The iterative methods to solve radio power control problems are a special case of parallel and distributed computation theory. There is a rich line of research on distributed iterative methods, with the corner stones [6, 15] by Tsitsiklis and Bertsekas. Most of the recent work focuses on convex problems, where duality and decomposition techniques can be used to distribute the computations over the involved nodes or agents of the network (see, e.g., [16] for a discussion of different methods). A framework that recently has attracted substantial attention is the Alternating method of Lagrangian Multipliers (ADMM) [17], which has been particularly successful for optimization problems in learning theory with huge data sets. These methods require a central entity that coordinates the nodes and their optimization subproblems. The problems are distributed in a computational sense, meaning that every node makes the computations coordinate by some central entity, rather than decentralized from an organizational point of view.

Decentralized solution methods are addressed by consensus methods, where all nodes are peers and compute the solution of network optimization problems by exchanging information with their local neighbors (see, e.g., [18–20]). In [20], each node has a local cost function and a local constraint set, both of which are assumed convex. The local problems are coupled through a common variable, and the global objective is to minimize the sum of all local costs. These powerful optimization methods are not easy to apply to the optimization of wireless networks, due to the slowness of the convergence of the message passing procedure. For example, consensus methods may converge with some hundreds of message exchanges that would take more time than the time needed to compute the optimal transmit power compared to the coherence time of the wireless channel. This means that when the solution would be computed by consensus methods, it is outdated.

In power control problems we often have in wireless networks, the global cost is not necessarily separable, nor convex, and each node controls only a subset of the variables. This is coherent with the lines of Yates’ framework [1] and the algorithms that are standard (e.g., [3–5, 7]). In fact, this paper will show that standard algorithms are encompassed in the Fast-Lipschitz framework. In [8], Yates’ framework is generalized to cover also some discrete implementations (e.g., [9], where the power updates are of a fixed size). Extensions of Yates’ framework have been proposed by Sung and Leung [10]. They consider opportunistic algorithms that are not standard by Yates’ definition. Instead, they introduce *type-II standard* functions and the more general *two-sided scalable* interference functions. These functions are shown to have the same fixed point properties as Yates’ standard functions, i.e., problem (3) with type-II standard constraints can be solved through repeated iterations of the constraints.

Further extension of the interference function framework have been proposed in [21, 22]. Specifically, [21] considers standard functions with a small modification, where the *scalability* property (see section (IV.B)) is replaced by the *scale invariance*, i.e., $\mathcal{I}(c\mathbf{p}) = c\mathcal{I}(\mathbf{p})$. It is shown how these functions can be interpreted as level curves of closed comprehensive sets. [22] replaces the scalability assumption of the standard framework by a requirement of weighted max-norm contraction. This allows to derive statements on the fixed points and convergence speed of iteration (4). However, neither of these extensions are concerned with notions of optimality of problems in the form of (3).

The algorithms above assume perfect knowledge of the interference functions. In [11] the convergence when some of the measurements required to evaluate the interference functions (e.g., the SNR samples) are stochastic and noisy or inaccurate is studied. [12] shows the

convergence of a general class of stochastic power control algorithms, given a suitable set of power update damping step lengths, and choice of these lengths are decentralized and improved in [13].

A different approach to power control is based on game theory, e.g., [2, 23–26]. [23] introduced an economic framework and modeled the power control problem as a non-cooperative game, where each mobile (node of the network) selfishly tries to maximize its local utility. The corresponding power updates (4) are then the best response of each mobile, and the fixed point corresponds to a Nash equilibrium. [24] investigates how this Nash equilibrium is affected by pricing of transmit powers. The game theoretic framework is flexible enough to model also cognitive radio networks with primary and secondary users [26]. An open problem in this line of research is that the resulting Nash equilibria typically do not correspond to a social optimum, meaning that there are other power configurations where all users are better off, or some global utility (such as total throughput) is higher.

In the power control algorithms mentioned above, the focus is not about the optimality of radio power control, but the convergence and existence of equilibria in distributed power updatings. In this paper, we consider the recently proposed theory of Fast-Lipschitz optimization to establish the optimality of power control algorithms. In particular, Fast-Lipschitz optimization is related to other techniques that replace the most common assumption of convexity with other conditions. Examples include the framework of monotonic optimization (e.g., [27–29]), where monotonic properties of the objective and the constraints allow for efficient solutions to the problem, or the framework of abstract optimization [30, 31] that generalizes linear programming in the sense that problems are solved by determining the subsets of the constraints that define the solution.

B. Notation

Vectors and matrices are denoted by bold lower and upper case letters, respectively. The components of a vector \mathbf{x} are denoted x_i or $[\mathbf{x}]_i$. Similarly, the elements of the matrix \mathbf{A} are denoted A_{ij} or $[\mathbf{A}]_{ij}$. The transpose of a vector or matrix is denoted \cdot^T . \mathbf{I} and $\mathbf{1}$ denote the identity matrix and the vector of all ones. A vector or matrix where all elements are zero is denoted by $\mathbf{0}$.

The gradient of a function is defined as $[\nabla \mathbf{f}(\mathbf{x})]_{ij} = \partial f_j(\mathbf{x}) / \partial x_i$, whereas $\nabla_i \mathbf{f}(\mathbf{x})$ denotes the i th row of $\nabla \mathbf{f}(\mathbf{x})$. Note that $\nabla \mathbf{f}(\mathbf{x})^k = (\nabla \mathbf{f}(\mathbf{x}))^k$, which has not to be confused with the k th derivative. The spectral radius is denoted $\rho(\cdot)$. Vector norms are denoted $\|\cdot\|$ and matrix

norms are denoted $\|\cdot\|$. Unless specified $\|\cdot\|$ and $\|\cdot\|$ denote arbitrary norms. $\|\mathbf{A}\|_\infty = \max_i \sum_j |A_{ij}|$ is the norm induced by the ℓ_∞ vector norm, where $\|\mathbf{x}\|_{\ell_\infty} = \max_i |x_i|$. These matrix norm definitions are coherent with [32].

All inequalities in this paper are intended *element-wise*, i.e., $\mathbf{A} \geq \mathbf{B}$ means $A_{ij} \geq B_{ij}$ for all i, j . We will also use the element-wise natural logarithm $\ln \mathbf{x} = [\ln x_1, \dots, \ln x_n]^T$ and the element-wise exponential $e^{\mathbf{x}} = [e^{x_1}, \dots, e^{x_n}]^T$.

Remark. The notation \mathcal{I} for interference functions does not follow the notational assumptions above. However, we will keep the notation to harmonize with existing literature [1, 4, 5, 21].

II. PROBLEM FORMULATION

We investigate a general form of the power minimization problem mentioned in the introduction section and having the general form

$$\begin{aligned} \min_{\mathbf{p}} \quad & \kappa(\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{p} \geq \mathcal{I}(\mathbf{p}). \end{aligned} \tag{5}$$

Throughout the paper we assume that $\kappa(\mathbf{p})$ and $\mathcal{I}(\mathbf{p})$ are differentiable. The cost function $\kappa(\mathbf{p})$ can be scalar or vector valued. Examples are $\kappa(\mathbf{p}) = \mathbf{p}$ or $\kappa(\mathbf{p}) = \mathbf{p}^T \mathbf{1}$. In practice, the powers must be positive and there is a maximum power that each transmitter can generate. Therefore, we will implicitly assume that there are the natural constraints $\mathbf{p} \in \mathcal{D}_{\mathbf{p}} = \{\mathbf{p} : \mathbf{p}_{\min} \leq \mathbf{p} \leq \mathbf{p}_{\max}\}$, where $\mathbf{p}_{\min} \geq \mathbf{0}$ and \mathbf{p}_{\max} are given constants.

The main problem this paper is concerned with, is when the iterations

$$\mathbf{p}^{k+1} := \mathcal{I}(\mathbf{p}^k) \tag{6}$$

solve optimization problem (5). We show under which conditions the general power control problem (5) is Fast-Lipschitz, which will allow us to establish the optimality of iterations (6). In particular, when problem (5) is Fast-Lipschitz, then the iterations (6) will converge to $\mathbf{p}^* = \mathcal{I}(\mathbf{p}^*)$ and \mathbf{p}^* is optimal for problem (5).

Remark 1. Note that the formulation of the iterations (6) is synchronous, i.e., every node must finish the computations and communications of round k before the next round $k+1$ can start. The algorithm we consider also converges asynchronously, under the assumption of bounded delays [6, 7, 10, 14]. However, since convergence properties are not the main focus of this paper, we restrict ourselves to the less cumbersome synchronous notation of (6).

III. FAST-LIPSCHITZ OPTIMIZATION

For the sake of self containment and the need for introducing a preliminary result, we now give a brief formal definition of Fast-Lipschitz problems. For a thorough discussion of Fast-Lipschitz properties we refer the reader to [14?].

Definition 2. A problem is said to be on *Fast-Lipschitz form* if it can be written as

$$\begin{aligned} \max \quad & \mathbf{f}_0(\mathbf{x}) \\ \text{s.t.} \quad & x_i \leq f_i(\mathbf{x}) \quad \forall i \in \mathcal{A} \\ & x_i = f_i(\mathbf{x}) \quad \forall i \in \mathcal{B}, \end{aligned} \tag{7}$$

where

- $\mathbf{f}_0 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a differentiable scalar ($m = 1$) or vector valued ($m \geq 2$) function.
- \mathcal{A} and \mathcal{B} are complementary subsets of $\{1, \dots, n\}$.
- For all i , $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is a differentiable function.

From the individual constraint functions we form the vector valued function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ as $\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) & \dots & f_n(\mathbf{x}) \end{bmatrix}^T$.

Remark 3. For the rest of the paper, we will restrict our attention to a *bounding box* $\mathcal{D} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\}$. We assume \mathcal{D} contains all candidates for optimality and that \mathbf{f} maps \mathcal{D} into \mathcal{D} , $\mathbf{f} : \mathcal{D} \rightarrow \mathcal{D}$. This box arises naturally in practice, since any real-world decision variable must be bounded.

Definition 4. A problem is said to be *Fast-Lipschitz* when it can be written on Fast-Lipschitz form and admits a unique Pareto optimal solution \mathbf{x}^* , defined as the unique solution to the system of equations

$$\mathbf{x}^* = \mathbf{f}(\mathbf{x}^*).$$

A problem written on Fast-Lipschitz form is not automatically Fast-Lipschitz. The appendix provides qualifying conditions which, when fulfilled, guarantee that problem (7) is Fast-Lipschitz (see Table 1 and Theorem 17).

The framework of Fast-Lipschitz optimization is formulated for maximization problems. Through a change of variables, any minimization problem can be formulated as an equivalent maximization problem. This is useful when dealing with power control problems, which are normally written by minimization. The following lemma shows how minimization problems fit into the Fast-Lipschitz framework.

Lemma 5 (Fast-Lipschitz minimization). *Consider*

$$\begin{aligned} \min \quad & \mathbf{g}_0(\mathbf{x}) \\ \text{s.t.} \quad & x_i \geq g_i(\mathbf{x}) \quad \forall i \in \mathcal{A} \\ & x_i = g_i(\mathbf{x}) \quad \forall i \in \mathcal{B}, \end{aligned} \tag{8}$$

where $\mathbf{g}_0(\mathbf{x})$, $\mathbf{g}(\mathbf{x}) = [g_i(\mathbf{x})]$, \mathcal{A} and \mathcal{B} fulfill the assumptions of Definition 2. Then, problem (8) is Fast-Lipschitz if $\mathbf{g}_0(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ fulfill the qualifying conditions.

Proof: Let $\mathbf{x} = -\mathbf{y}$ and form the equivalent problem

$$\begin{aligned} \max \quad & -\mathbf{g}_0(-\mathbf{y}) = \mathbf{f}_0(\mathbf{y}) \\ \text{s.t.} \quad & y_i \leq -g_i(-\mathbf{y}) = f_i(\mathbf{y}) \quad \forall i \in \mathcal{A} \\ & y_i = -g_i(-\mathbf{y}) = f_i(\mathbf{y}) \quad \forall i \in \mathcal{B}. \end{aligned} \tag{9}$$

In order to check the qualifying conditions one needs $\nabla \mathbf{f}_0(\mathbf{y})$ and $\nabla \mathbf{f}(\mathbf{y})$. But

$$(\nabla_{\mathbf{y}} \mathbf{f}_0(\mathbf{y}))^T = \frac{\partial \mathbf{f}_0(\mathbf{y})}{\partial \mathbf{y}} = \frac{\partial (-\mathbf{g}_0(\mathbf{x}))}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{y}} = -\frac{\partial \mathbf{g}_0(\mathbf{x})}{\partial \mathbf{x}} (-1) = (\nabla_{\mathbf{x}} \mathbf{g}_0(\mathbf{x}))^T,$$

and analogously, $\nabla_{\mathbf{y}} \mathbf{f}(\mathbf{y}) = \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x})$. Since $\mathbf{g}_0(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ fulfill the qualifying conditions, the equivalent problem (9) is Fast-Lipschitz. \square

We are now in the position of introducing the core contribution of the paper in the following section.

IV. TWO-SIDED SCALABLE PROBLEMS AND FAST-LIPSCHITZ OPTIMIZATION

In this section we examine the relations between standard functions (in IV.B), type-II standard functions (in IV.C), and Fast-Lipschitz optimization. This will allow us to establish the core results that: 1) all standard power control problems are Fast Lipschitz; 2) there exist non standard interference functions whose fixed point is the optimal of a power control problem; 3) type II power control iterations are optimal within some conditions. We begin by briefly recalling the concept of two-sided scalability, which we will use to put standard and type-II standard functions in the Fast-Lipschitz framework.

A. Preliminary results on two-sided scalable functions

This subsection presents preliminary results for the upcoming sections on standard and type-II standard problems. The main result of this section is Lemma 8, which will allow us to establish the contractivity of standard and type-II standard functions. Contractivity is one of the main ingredients for Fast-Lipschitz optimization.

Definition 6 ([10]). A function $\mathcal{I}(\mathbf{p})$ is *two-sided scalable* if for all $c > 1$ and all $(1/c)\mathbf{p} \leq \mathbf{q} \leq c\mathbf{p}$, it holds that

$$(1/c)\mathcal{I}(\mathbf{p}) < \mathcal{I}(\mathbf{q}) < c\mathcal{I}(\mathbf{p}). \quad (10)$$

Proposition 7 ([10, Prop. 4]). *If a function is either standard or type-II standard, then it is also two-sided scalable.*

The following lemma shows that two-sided scalable functions are shrinking maps with gradients of one-norm less than one. The lemma is based on the logarithmic transformations proposed in [33]. It will be used in the main results of sections IV.B and IV.C.

Lemma 8. *Let $\mathbf{x} \triangleq \ln \mathbf{p}$ and $\mathbf{f}(\mathbf{x}) \triangleq \ln \mathcal{I}(\mathbf{x})$. If $\mathcal{I}(\mathbf{p})$ is two-sided scalable, then*

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\|_\infty < \|\mathbf{x} - \mathbf{y}\|_\infty$$

for all \mathbf{x}, \mathbf{y} , and

$$\|\nabla \mathbf{f}(\mathbf{x})\|_1 < 1.$$

Proof: By [10, Lemma 7], two-sided scalability implies

$$\max_i \left\{ \max \left\{ \frac{\mathcal{I}_i(\mathbf{x})}{\mathcal{I}_i(\mathbf{y})}, \frac{\mathcal{I}_i(\mathbf{y})}{\mathcal{I}_i(\mathbf{x})} \right\} \right\} < \max_i \left\{ \max \left\{ \frac{x_i}{y_i}, \frac{y_i}{x_i} \right\} \right\}$$

for all i . Since the logarithm is strictly increasing, this is equivalent to

$$\max_i \left\{ \max \left\{ \ln \left(\frac{\mathcal{I}_i(\mathbf{p})}{\mathcal{I}_i(\mathbf{q})} \right), \ln \left(\frac{\mathcal{I}_i(\mathbf{q})}{\mathcal{I}_i(\mathbf{p})} \right) \right\} \right\} < \max_i \left\{ \max \left\{ \ln \left(\frac{p_i}{q_i} \right), \ln \left(\frac{q_i}{p_i} \right) \right\} \right\}.$$

Inserting $\mathbf{p} = e^{\mathbf{x}}$ and $\mathbf{q} = e^{\mathbf{y}}$ gives

$$\begin{aligned} \max_i \{ |\ln \mathcal{I}_i(e^{\mathbf{x}}) - \ln \mathcal{I}_i(e^{\mathbf{y}})| \} &< \max_i \{ |\ln e^{x_i} - \ln e^{y_i}| \}, \\ \Leftrightarrow |f_i(\mathbf{x}) - f_i(\mathbf{y})| &< |x_i - y_i|. \end{aligned}$$

Since this holds for all components i , we have

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\|_\infty < \|\mathbf{x} - \mathbf{y}\|_\infty$$

by definition of norm infinity.

Denote $\mathbf{v} = \arg \max_{\|\mathbf{u}\|_\infty=1} \|\nabla \mathbf{f}(\mathbf{x})^T \mathbf{u}\|_\infty$. By definition, we have $\|\mathbf{v}\|_\infty = 1$ and

$$\|\nabla \mathbf{f}(\mathbf{x})^T \mathbf{v}\|_\infty = \left\| \nabla \mathbf{f}(\mathbf{x})^T \right\|_\infty = \|\nabla \mathbf{f}(\mathbf{x})\|_1.$$

By defining $\mathbf{y} = \mathbf{x} + \epsilon \mathbf{v}$, with ϵ positive scalar, we have

$$1 > \frac{\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\|_\infty}{\|\mathbf{y} - \mathbf{x}\|_\infty} = \frac{\|\mathbf{f}(\mathbf{x} + \epsilon \mathbf{v}) - \mathbf{f}(\mathbf{x})\|_\infty}{\|\epsilon \mathbf{v}\|_\infty} = \left\| \frac{\mathbf{f}(\mathbf{x} + \epsilon \mathbf{v}) - \mathbf{f}(\mathbf{x})}{\epsilon} \right\|_\infty,$$

and in the limit $\epsilon \rightarrow 0$,

$$1 > \lim_{\epsilon \rightarrow 0} \frac{\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\|_\infty}{\|\mathbf{y} - \mathbf{x}\|_\infty} = \lim_{\epsilon \rightarrow 0} \left\| \frac{\mathbf{f}(\mathbf{x} + \epsilon \mathbf{v}) - \mathbf{f}(\mathbf{x})}{\epsilon} \right\|_\infty = \left\| \nabla \mathbf{f}(\mathbf{x})^T \mathbf{v} \right\|_\infty = \|\nabla \mathbf{f}(\mathbf{x})\|_1.$$

This concludes the proof. \square

The result shows that two-sided scalable functions are shrinking maps, but it does not establish the amount of slack in $\|\nabla \mathbf{f}(\mathbf{x})\|_1 < 1$. This slack is useful for several reasons. The first reason is that any amount of slack makes the function $\mathbf{f}(\mathbf{x})$ a contraction, thereby guaranteeing a unique fixed point. This is assumed in the Fast-Lipschitz qualifying conditions, e.g., in (GQC.b). The lack of knowledge on the amount of slack is not a problem in practice, since the bounds $\mathbf{p}_{\min} \leq \mathbf{p} \leq \mathbf{p}_{\max}$ form a closed bounded region of \mathfrak{R}^n . Therefore, $\|\nabla \mathbf{f}(\mathbf{x})\|_1$ attains a minimum (call this value α) in that region, so $\|\nabla \mathbf{f}(\mathbf{x})\|_1 \leq \alpha < 1$ for those \mathbf{x} that are of interest. Secondly (and more importantly), the qualifying conditions other than \mathbf{Q}_1 typically require $\|\nabla \mathbf{f}(\mathbf{x})\|_\infty < c$ for some $c \in (0, 1]$. Lemma 8 is therefore only of use if $c = 1$. This special case is exploited in Section IV.C. Now, based on this lemma, we are in the position to show that power control problems (5) with standard interference functions are a special case and Fast-Lipschitz optimization.

B. Standard functions and Fast-Lipschitz optimization

In this section we recall Yates' framework of standard functions and show that a problem (5) with standard interference function constraints has an equivalent problem formulation that is Fast-Lipschitz.

Definition 9 ([1]). The function $\mathcal{I}(\mathbf{p})$ is *standard* if for all $\mathbf{p}, \mathbf{q} \geq \mathbf{0}$, the following properties are satisfied.

$$\text{Positivity:} \quad \mathcal{I}(\mathbf{p}) > \mathbf{0} \tag{11a}$$

$$\text{Monotonicity:} \quad \mathbf{p} \geq \mathbf{q} \Rightarrow \mathcal{I}(\mathbf{p}) \geq \mathcal{I}(\mathbf{q}) \tag{11b}$$

$$\text{Scalability:} \quad c > 1 \Rightarrow \mathcal{I}(c\mathbf{p}) < c\mathcal{I}(\mathbf{p}) \tag{11c}$$

The monotonicity requirement (11b) can equivalently be formulated $\nabla \mathcal{I}(\mathbf{p}) \geq \mathbf{0}$ for all $\mathbf{p} \geq \mathbf{0}$.

There is a relation between standard functions and two-sided scalable functions. Note that (10) multiplied by a positive scalar $c > 0$ implies $(c^2 - 1)\mathcal{I}(\mathbf{p}) > \mathbf{0}$, so any two-sided

scalable function must also be positive, i.e., $\mathcal{I}(\mathbf{p}) > 0$ [10, Lemma 6]. If \mathbf{q} in Definition 6 is chosen as $\mathbf{q} = c\mathbf{p}$, inequalities (10) become

$$(1/c)\mathcal{I}(\mathbf{p}) < \mathcal{I}(c\mathbf{p}) < c\mathcal{I}(\mathbf{p}), \quad (12)$$

so a two-sided scalable function is always scalable (11c). Any two-sided scalable function $\mathcal{I}(\mathbf{p})$ is therefore standard if $\nabla \mathcal{I}(\mathbf{p}) \geq 0$.

Proposition 10 ([1]). *Assume that the power optimization problem (3) is feasible. Then, the standard interference function $\mathcal{I}(\mathbf{p})$ has a unique fixed point \mathbf{p}^* that is the solution to (3).*

We now show that problem (3) with standard interference constraints fall under the Fast-Lipschitz framework. To this end we consider problem (5), with the general cost function $\kappa(\mathbf{p})$, which we assume be differentiable. All Fast-Lipschitz qualifying conditions require that the norm of the constraint function gradient be small enough. We will show this through Lemma 8, wherefore we investigate problem (5) after a change of variables.

To this end, we let $\mathbf{x} \triangleq \ln \mathbf{p}$ as the logarithm of the power variables. This gives $\mathbf{p} \triangleq e^{\mathbf{x}}$ and the equivalent problems

$$\begin{aligned} \min_{\mathbf{x}} \quad & \kappa(e^{\mathbf{x}}) \\ \text{s.t.} \quad & e^{\mathbf{x}} = \mathcal{I}(e^{\mathbf{x}}) \end{aligned}$$

and, because the logarithm is strictly increasing,

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{f}_0(\mathbf{x}) \triangleq \kappa(e^{\mathbf{x}}) \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{f}(\mathbf{x}) \triangleq \ln \mathcal{I}(e^{\mathbf{x}}). \end{aligned} \quad (13)$$

If problem (13) is Fast-Lipschitz, then $\mathbf{x}^* = \mathbf{f}(\mathbf{x}^*)$ is the unique Pareto optimal point for (13), whereby

$$e^{\mathbf{x}^*} = \mathbf{p}^* = \mathcal{I}(\mathbf{p}^*)$$

is optimal for problem (5).

The following result shows how power control problems with standard constraint functions, if differentiable, have an equivalent Fast-Lipschitz problem formulation.

Theorem 11. *Consider problem (5) and let $\mathcal{I}(\mathbf{p})$ be differentiable and standard. If $\nabla \kappa(\mathbf{p}) \geq 0$ with non-zero rows, then the equivalent problem (13) is Fast-Lipschitz and $\mathbf{p}^* = \mathcal{I}(\mathbf{p}^*)$ is optimal in problem (5).*

Proof: We show that problem (13) is Fast-Lipschitz by qualifying condition Q_1 . The gradients of problem (13) are given by

$$\nabla f_0(\mathbf{x}) = \text{diag}(\mathbf{p}) \nabla \kappa(\mathbf{p}), \quad (14a)$$

$$\nabla f(\mathbf{x}) = \text{diag}(\mathbf{p}) \nabla \mathcal{I}(\mathbf{p}) \text{diag}(1/\mathcal{I}(\mathbf{p})), \quad (14b)$$

where $\mathbf{p} = e^{\mathbf{x}}$ and $[\text{diag}(1/\mathcal{I}(\mathbf{p}))]_{ii} = 1/\mathcal{I}_i(\mathbf{p})$. Since $\mathbf{p} = e^{\mathbf{x}} \geq \mathbf{0}$, $\mathcal{I}(\mathbf{p}) > \mathbf{0}$ and $\nabla \mathcal{I}(\mathbf{p}) \geq \mathbf{0}$, the gradients (14) fulfill $\nabla f(\mathbf{x}) \geq \mathbf{0}$ and $\nabla f_0(\mathbf{x}) \geq \mathbf{0}$ with non-zero rows (these are conditions $(Q_1.c)$ and $(Q_1.a)$ respectively). If $\mathcal{I}(\mathbf{p})$ is standard, it is also two-sided scalable by Proposition 7, so $\|\nabla f(\mathbf{x})\|_1 < 1$ by Lemma 8 (condition $(Q_1.b)$). Problem (13) is therefore Fast-Lipschitz by qualifying condition Q_1 , and $\mathbf{x}^* = \mathbf{f}(\mathbf{x}^*)$. Taking the exponential of the previous relation gives $\mathbf{p}^* = \mathcal{I}(\mathbf{p}^*)$. This concludes the proof. \square

While Proposition 10 states that the fixed point of standard constraints minimize the powers in a Pareto sense, i.e., $\kappa(\mathbf{p}) = \mathbf{p}$, Theorem 11 accepts any non-decreasing $\kappa(\mathbf{p})$. The requirement that $\nabla \kappa(\mathbf{p})$ have non-zero rows simply means that each variable p_i has an effect on at least one component of the cost at each \mathbf{p} . For scalar values cost functions, this is the same as requiring κ to be strictly increasing, $\nabla \kappa(\mathbf{p}) > \mathbf{0}$.

Theorem 11 is not a generalization of Theorem 10 in practice, since a minimization of \mathbf{p} is equivalent to a minimization of an increasing function of \mathbf{p} . The novelty here is instead that standard problems falls within the broader class of Fast-Lipschitz problems. Therefore, we can have non standard interference functions that can lead to optimally by distributed iterative power control algorithms. In the next subsection we will continue to show how type-II standard functions relates to Fast-Lipschitz optimization.

C. Type-II standard functions and Fast-Lipschitz optimization

As the standard functions are monotonically increasing, transmit nodes following (6) will always increase their power when their transmission environment is worsened by higher interference. A receiver node experiencing a deep fade will therefore need a very high transmit power, thereby increasing interference for the other receiver nodes in the network. This is not a good strategy, for example, in delay tolerant applications, where transmit nodes can adjust their transmission rates and higher throughput can be achieved by prioritizing receiver nodes experiencing low interference. One such strategy is to keep the signal-to-interference product constant, which results in update functions (6) that are monotonically decreasing,

and therefore not standard. This is addressed in [10], where Sung and Leung extends Yates' framework with type-II standard functions.

Definition 12 ([10]). The function $\mathcal{I}(\mathbf{p})$ is *type-II standard* if for all $\mathbf{p}, \mathbf{q} \geq \mathbf{0}$, the following properties are satisfied:

$$\text{Type-II Monotonicity: } \mathbf{p} \leq \mathbf{q} \Rightarrow \mathcal{I}(\mathbf{p}) \geq \mathcal{I}(\mathbf{q}) \quad (15a)$$

$$\text{Type-II Scalability: } c > 1 \Rightarrow \mathcal{I}(c\mathbf{p}) > (1/c)\mathcal{I}(\mathbf{p}) \quad (15b)$$

As in the case of standard functions, the monotonicity property (15a) can be written $\nabla \mathcal{I}(\mathbf{p}) \leq \mathbf{0}$ for all \mathbf{p} . Note also from (12) that all two-sided scalable functions are type-II scalable (15b) so any two-sided scalable function $\mathcal{I}(\mathbf{p})$ where $\nabla \mathcal{I}(\mathbf{p}) \leq \mathbf{0}$ is also type-II standard. Type-II standard functions converge in the same way as standard functions - if a fixed point \mathbf{p}^* exists, then iteration (6) converges to \mathbf{p}^* [10, Thm. 3].

When considering opportunistic algorithms, $\mathcal{I}_i(\mathbf{p})$ no longer has the interpretation of “interference that receiver node i must overpower asking transmit node i to use a power p_i high enough”. There are no longer any explicit constraints $\mathbf{p} \geq \mathcal{I}(\mathbf{p})$ underlying the algorithm, and \mathcal{I} might not even have a physical meaning. The framework of two-scalable functions guarantees that the iterations (6) converge to a fix point also in the case of type-II standard interference functions, but the optimality meaning of this fixed point is no longer clear. Therefore, in the following we consider a function \mathcal{I} of type-II and assume it comes from a problem of the form (5). With the framework of Fast-Lipschitz optimization we characterize type-II standard power control problems to show that the fixed point is also optimal for optimization in the form (5). This is an important result that we can establish by Fast-Lipschitz optimization.

As in the of standard functions in Section (IV.B), we examine the problem in logarithmic variables $\mathbf{x} = \ln \mathbf{p}$ and arrive at the equivalent problem (13), with gradients given by (14).

Theorem 13. Assume $\mathcal{I}(\mathbf{p})$ be differentiable and type-II standard, and consider $\mathbf{f}(\mathbf{x}) = \ln \mathcal{I}(e^{\mathbf{x}})$. Let $\mathbf{B} = [B_{ij}]$ such that

$$B_{ij} = \max_{\mathbf{x}} |\nabla_i f_j(\mathbf{x})| = \max_{\mathbf{p}} \left| \nabla_i \mathcal{I}_j(\mathbf{p}) \frac{p_i}{\mathcal{I}_j(\mathbf{p})} \right|,$$

and assume $\rho(\mathbf{B}) < 1$. Let $\mathbf{c} > \mathbf{0}$ be an arbitrary (positive) vector in \mathbb{R}^n and let $\mathbf{h}(z) \in \mathbb{R}^m$ be any strictly increasing function of one variable. Then, problem (5) is Fast-Lipschitz if

$$\mathbf{s} = (\mathbf{I} - \mathbf{B})^{-1} \mathbf{c} \quad (16)$$

and

$$\kappa(\mathbf{p}) = \mathbf{h}(\prod_i p_i^{s_i}). \quad (17)$$

Proof: Since $\rho(\mathbf{B}) < 1$, $(\mathbf{I} - \mathbf{B})^{-1}$ is invertible and

$$\mathbf{s} = (\mathbf{I} - \mathbf{B})^{-1} \mathbf{c} = \sum_{k=0}^{\infty} \mathbf{B}^k \mathbf{c} > \mathbf{0}.$$

Let $\mathbf{S} = \text{diag}(\mathbf{s})$ and introduce the scaled variables $\mathbf{y} = \mathbf{S}\mathbf{x}$. The inverse of \mathbf{S} exists and is positive since $\mathbf{s} > \mathbf{0}$. Inserting $\mathbf{x} = \mathbf{S}^{-1}\mathbf{y}$ in (13) gives the equivalent problem

$$\begin{aligned} \min_{\mathbf{y}} \quad & \mathbf{g}_0(\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{y} \geq \mathbf{g}(\mathbf{y}) = \mathbf{S} \mathbf{f}(\mathbf{S}^{-1}\mathbf{y}), \end{aligned} \quad (18)$$

where

$$\mathbf{g}_0(\mathbf{y}) = \mathbf{f}_0(\mathbf{S}^{-1}\mathbf{y}) = \mathbf{h}(\prod_i e^{y_i}) = \mathbf{h}(e^{\sum_i y_i}) := \mathbf{h}(z(\mathbf{y})).$$

The gradients of the problem are

$$[\nabla \mathbf{g}_0(\mathbf{y})]_{ij} = h'_j(z(\mathbf{y})) \frac{\partial z(\mathbf{y})}{\partial y_i} = h'_j(z(\mathbf{y})) z(\mathbf{y})$$

and $\nabla \mathbf{g}(\mathbf{y}) = \mathbf{S}^{-1} \nabla \mathbf{f}(\mathbf{x}) \mathbf{S}$, where $\mathbf{x} = \mathbf{S}^{-1}\mathbf{y}$.

We will show that problem (18) is Fast-Lipschitz by \mathbf{Q}_2 . Condition $(\mathbf{Q}_2.b)$ is fulfilled because $\mathcal{I}(\mathbf{p}) > \mathbf{0}$ and $\nabla \mathcal{I}(\mathbf{p}) \leq \mathbf{0}$ when \mathcal{I} is type-II standard. This gives

$$\nabla_i g_j(\mathbf{y}) = \nabla_i f_j(\mathbf{x}) \frac{s_j}{s_i} = \left(\nabla_i \mathcal{I}_j(\mathbf{p}) \frac{p_i}{\mathcal{I}_j(\mathbf{p})} \right) \frac{s_j}{s_i} \leq 0 \quad \forall i, j,$$

since $\mathbf{s} > \mathbf{0}$ and $\mathbf{p} = e^{\mathbf{x}} = e^{\mathbf{S}^{-1}\mathbf{y}} \geq \mathbf{0}$.

Condition $(\mathbf{Q}_2.c)$ requires

$$\|\nabla \mathbf{g}(\mathbf{y})\|_{\infty} < \min_j \frac{\min_i [\nabla \mathbf{g}_0(\mathbf{y})]_{ij}}{\max_i [\nabla \mathbf{g}_0(\mathbf{y})]_{ij}}, \quad (19)$$

which is true by construction. To see this, note that

$$\min_j \frac{\min_i [\nabla \mathbf{g}_0(\mathbf{y})]_{ij}}{\max_i [\nabla \mathbf{g}_0(\mathbf{y})]_{ij}} = \min_j \frac{\min_i h'_j(z(\mathbf{y})) z(\mathbf{y})}{\max_i h'_j(z(\mathbf{y})) z(\mathbf{y})} = 1.$$

The left side of (19) therefore requires

$$\|\nabla \mathbf{g}(\mathbf{y})\|_{\infty} = \|\mathbf{S}^{-1} \nabla \mathbf{f}(\mathbf{x}) \mathbf{S}\|_{\infty} = \max_i \sum_j \left| \nabla_i f_j(\mathbf{x}) \frac{s_j}{s_i} \right| < 1$$

for all \mathbf{y} and $\mathbf{x} = \mathbf{S}^{-1}\mathbf{y}$. Since $s_i > 0$ and $|\nabla_i f_j(\mathbf{x})| \leq B_{ij}$ for all \mathbf{x} , this holds if

$$\max_i \sum_j B_{ij} \frac{s_j}{s_i} < 1,$$

or equivalently, if $\sum_j B_{ij}s_j < s_i$ for all i . This is the i th row of $(\mathbf{I} - \mathbf{B})\mathbf{s} > \mathbf{0}$, which holds by construction since

$$(\mathbf{I} - \mathbf{B})\mathbf{s} = (\mathbf{I} - \mathbf{B})(\mathbf{I} - \mathbf{B})^{-1}\mathbf{c} = \mathbf{c} > \mathbf{0}.$$

Finally, condition (Q₂.a) is easily checked because

$$[\nabla \mathbf{g}_0(\mathbf{y})]_{ij} = h'_j(z(\mathbf{y}))z(\mathbf{y}) > \mathbf{0},$$

which holds due to that $h_j(z)$ is increasing and $z(\mathbf{y}) = e^{\sum_i y_i} > 0$. This concludes the proof. \square

The form $\kappa(\mathbf{p}) = \mathbf{h}(z(\mathbf{p}))$, with \mathbf{h} being an increasing function, implies that all cost functions κ that can be handled with Theorem 13 are equivalent to the scalar cost $\kappa_0(\mathbf{p}) = \prod_i p_i^{s_i}$ obtained when $\mathbf{h}(z) = z$. If one instead chooses $\mathbf{h}(z) = \ln z$ the cost becomes $\kappa(\mathbf{p}) = \mathbf{s}^T \ln \mathbf{p}$, i.e., a weighted sum of the power logarithms. Equation (16) states that weighting \mathbf{s} should lie in the interior of the cone spanned by the columns of $(\mathbf{I} - \mathbf{B}^T)^{-1}$.

The assumption $\rho(\mathbf{B}) < 1$ is crucial for Theorem 13 to hold, since it guarantees the existence of a positive scaling matrix \mathbf{S} . The assumptions assure that $\rho(\nabla \mathbf{f}(\mathbf{x})) \leq \|\nabla \mathbf{f}(\mathbf{x})\|_1 < 1$, so $\rho(\mathbf{B}) < 1$ surely holds if there is a point \mathbf{x}^B such that $\mathbf{B} = |\nabla \mathbf{f}(\mathbf{x}^B)|$. This means that all elements of $\nabla \mathbf{f}(\mathbf{x})$ are minimized at the common point \mathbf{x}^B . The simplest case where this is true is when $\nabla \mathbf{f}(\mathbf{x}) = \mathbf{A}^T$ is constant. This requires an $\mathcal{I}(\mathbf{p})$ of the form

$$\begin{aligned} \mathcal{I}_i(\mathbf{p}) &= \exp(\mathbf{A} \ln \mathbf{p} + \mathbf{b}) = e^{b_i} \exp\left(\sum_j A_{ij} \ln p_j\right) = e^{b_i} \exp\left(\sum_j \ln p_j^{A_{ij}}\right) \\ &= e^{b_i} \exp\left(\ln \prod_j p_j^{A_{ij}}\right) = e^{b_i} \prod_j p_j^{A_{ij}}, \end{aligned}$$

i.e., $\mathcal{I}_i(\mathbf{p})$ should be a monomial. If problem (5) has the basic cost function

$$\kappa_0(\mathbf{p}) = \prod_i p_i^{s_i}$$

from above, which also is a monomial, the problem is a geometric optimization problem [34]. Interestingly, geometric problems become convex with the change of variables $\mathbf{x} \triangleq \ln \mathbf{p}$, the same variable transformation used throughout this section.

From the discussion above, we can establish the following new qualifying condition:

Qualifying Condition 6	
\mathbf{Q}_6	(Q _{6.a}) $\nabla f(\mathbf{x})^2 \geq \mathbf{0}$, (e.g., $\nabla f(\mathbf{x}) \leq \mathbf{0}$)
	(Q _{6.b}) $ \nabla f(\mathbf{x}) \leq \mathbf{B}$ and $\rho(\mathbf{B}) < 1$
	(Q _{6.c}) $\mathbf{f}_0(\mathbf{x}) = \mathbf{h}(\sum_i s_i x_i)$ for a strictly increasing function $\mathbf{h}(z)$ where $\mathbf{s} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{c}$ and $\mathbf{c} > \mathbf{0}$

Observe that the notation in (Q_{6.b}) means the absolute value, not the norm. The new condition is numbered 6, although it is the 4th condition of this paper (see the appendix), but there are five known qualifying conditions in Fast-Lipschitz optimization [?], the last two are not used in this paper and therefore are not included in the appendix.

Theorem 14. *Assume problem (7) is feasible, and that qualifying condition \mathbf{Q}_6 above holds for every $\mathbf{x} \in \mathcal{D}$. Then, the problem is Fast-Lipschitz, i.e., the unique Pareto optimal solution is given by $\mathbf{x}^* = \mathbf{f}(\mathbf{x}^*)$.*

Proof: The proof is analogous to that of Theorem 13. □

The simplest form of the function $\mathbf{f}_0(\mathbf{x})$ in \mathbf{Q}_6 arises from $\mathbf{h}(z) = z$. This means that $\mathbf{f}_0(\mathbf{x}) = \mathbf{s}^T \mathbf{x}$ is a weighted sum of the power logarithms, where the weights \mathbf{s} are closely related to the constraint gradient.

In the following section, we turn our attention to a class of power control problems that do not have monotonic constraint functions.

V. ABSOLUTELY SUBHOMOGENEOUS INTERFERENCE FUNCTIONS

In the previous sections we examined two-sided scalable functions that were monotonically increasing (standard) and monotonically decreasing (type-II standard). In the following we give an example of a problem formulation where the constraints are not monotonic, hence neither standard, nor type-II standard. We show convergence and optimality through Fast-Lipschitz optimization, which was not known before.

The example builds upon the problem formulation in [35]. Once again we consider problem (5) and assume that the cost function $\kappa(\mathbf{p})$ is increasing in \mathbf{p} . The formulation in [35] starts with the affine SNR model (2), but adds a stochastic channel and outage as follows. Let

$$\mathcal{I}_i(\mathbf{p}) = \frac{\tau_i}{g_{ii}} \left(\sum_{j \neq i} g_{ij} p_j + \eta_i \right) \quad (20)$$

represent the *expected* power needed to reach the SNR target τ_i , and model the stochastic gain from transmitter i to receiver i by $g_{ii}\Theta_i$ where Θ_i is a stochastic variable describing the fading of the wireless channel. Furthermore, allow each transmitter to send only if the required power (to reach the SNR target) is lower than some bound b . Combining the two effects gives the new power control law

$$p_i^{k+1} = h\left(\frac{\mathcal{I}_i(\mathbf{p}^k)}{\Theta_i}\right), \quad (21)$$

where

$$h(x) = \begin{cases} x & \text{if } x \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

The fast timescale of the fading Θ_i makes it hard to track and measure in practice. Instead, let each transmitter node update its transmit power according to the expectation (21), i.e.,

$$p_i^{k+1} = \mathbb{E}_{\Theta_i} \left[h\left(\frac{\mathcal{I}_i(\mathbf{p}^k)}{\Theta_i}\right) \right] \triangleq \Phi_i(\mathcal{I}_i(\mathbf{p}^k)). \quad (22)$$

The expectation acts to smooth the discontinuous properties of $h(\cdot)$, and $\Phi_i(\mathcal{I}_i(\mathbf{p}))$ is called the smoothed interference function of node (or mobile equipment) i . The iterations in (22) can be seen as a possible solution algorithm for a power control problem of the type

$$\begin{aligned} \min \quad & \kappa(\mathbf{p}) \\ \text{s.t.} \quad & p_i \geq f_i(\mathbf{p}) \triangleq \Phi_i(\mathcal{I}_i(\mathbf{p})) \quad \forall i. \end{aligned} \quad (23)$$

However, the nature of $h(x)$ will make $f_i(\mathbf{p})$ non-monotonic, regardless of underlying assumptions on $\mathcal{I}_i(\mathbf{p})$. Therefore, neither the standard, nor the type-II standard interference function approach applies here. To study the convergence properties of iterations based on these functions, [35] introduces *absolutely subhomogeneous* functions, fulfilling

$$e^{-|a|}\Phi(\mathbf{x}) \leq \Phi(e^a\mathbf{x}) \leq e^{|a|}\Phi(\mathbf{x})$$

for every $\mathbf{x} \geq \mathbf{0}$ and all scalars a . Note that absolute subhomogeneity is implied by two-sided scalability. In [35] it is shown that, if for each i ,

- $\mathcal{I}_i(\mathbf{p})$ is *standard*, and
- $\Phi_i(x) = \mathbb{E}_{\Theta_i} [h(x/\Theta_i)]$ is bounded and absolutely subhomogeneous,

then the sequence (22) will converge to a fixed point. However, nothing is said in [35] about the optimality of this fixed point.

Our approach is to use Fast-Lipschitz optimization and qualifying condition Q_3 , which has no requirements on the monotonicity of $\mathbf{f}(\mathbf{x})$. Consider again problem (23). If

$$\mathbf{f}(\mathbf{p}) = [f_1(\mathbf{x}), \dots, f_n(\mathbf{x})]^T$$

and $\kappa(\mathbf{p})$ fulfill $(Q_3.b)$, i.e., if $\nabla \kappa(\mathbf{p}) > \mathbf{0}$ and

$$\|\nabla \mathbf{f}(\mathbf{p})\|_\infty < \frac{\mathbf{q}(\mathbf{p})}{1 + \mathbf{q}(\mathbf{p})},$$

where

$$\mathbf{q}(\mathbf{p}) = \min_j \frac{\min_i \nabla_i \kappa_j(\mathbf{p})}{\max_i \nabla_i \kappa_j(\mathbf{p})},$$

then problem (23) is Fast-Lipschitz and the iterations (22) will converge to the optimal solution of (23). In the previous sections, we used properties of standard and type-II standard functions to show that the gradient norm $\|\nabla \mathbf{f}\|_\infty$ was small enough. In this section, we obtain the bound directly from $\|\nabla \mathcal{I}\|_\infty$ by using the following result:

Lemma 15. *Let $\theta_j(y)$ be the pdf of the channel fading coefficient Θ_j , consider $z > 0$ and define*

$$\Omega_j(z) \triangleq \int_z^\infty \frac{\theta_j(y)}{y} dy - \theta_j(z). \quad (24)$$

Then, the infinity norm of the constraint function of problem (23) and the infinity norm of the underlying interference function $\mathcal{I}(\mathbf{p})$ in (20) fulfill

$$\|\nabla \mathbf{f}(\mathbf{p})\|_\infty \leq \max_i |\Omega_i(\mathcal{I}_i(\mathbf{p})/b)| \|\nabla \mathcal{I}(\mathbf{p})\|_\infty. \quad (25)$$

Proof: Dropping the explicit \mathbf{p} -dependence of \mathcal{I}_j and f_j , we have

$$\begin{aligned} f_j &= \mathbb{E}_{\Theta_j}[h(\mathcal{I}_j/\Theta_j)] = \int_0^\infty h(\mathcal{I}_j/y) \theta_j(y) dy \\ &= \int_0^{\mathcal{I}_j/b} \underbrace{h(\mathcal{I}_j/y)}_{=0} \theta_j(y) dy + \int_{\mathcal{I}_j/b}^\infty \underbrace{h(\mathcal{I}_j/y)}_{=\mathcal{I}_j/y} \theta_j(y) dy = \mathcal{I}_j \int_{\mathcal{I}_j/b}^\infty \frac{\theta_j(y)}{y} dy, \end{aligned}$$

because

$$h(\mathcal{I}_j/y) = \begin{cases} \mathcal{I}_j/y & \text{if } \mathcal{I}_j/y \leq b \iff y \geq \mathcal{I}_j/b, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\frac{df_j}{d\mathcal{I}_j} = \frac{d}{d\mathcal{I}_j} \left(\mathcal{I}_j \int_{\mathcal{I}_j/b}^\infty \frac{\theta_j(y)}{y} dy \right) = \int_{\mathcal{I}_j/b}^\infty \frac{\theta_j(y)}{y} dy - \theta(\mathcal{I}_j/b) \triangleq \Omega_j(\mathcal{I}_j/b).$$

Returning to full notation, we have

$$\frac{\partial f_j(\mathbf{p})}{\partial p_i} = \frac{df_j(\mathbf{p})}{d\mathcal{I}_j(\mathbf{p})} \frac{\partial \mathcal{I}_j(\mathbf{p})}{\partial p_i} = \Omega_j(\mathcal{I}_j(\mathbf{p})/b) \frac{\partial \mathcal{I}_j(\mathbf{p})}{\partial p_i}.$$

It follows that

$$\begin{aligned} \|\nabla \mathbf{f}(\mathbf{p})\|_\infty &= \max_i \sum_j \left| \frac{\partial f_j(\mathbf{p})}{\partial p_i} \right| = \max_i \sum_j \left| \Omega_j(\mathcal{I}_j(\mathbf{p})/b) \frac{\partial \mathcal{I}_j(\mathbf{p})}{\partial p_i} \right| \\ &\leq \max_j |\Omega_j(\mathcal{I}_j(\mathbf{p})/b)| \cdot \max_i \sum_j \left| \frac{\partial \mathcal{I}_j(\mathbf{p})}{\partial p_i} \right| \\ &= \max_j |\Omega_j(\mathcal{I}_j(\mathbf{p})/b)| \cdot \|\nabla \mathcal{I}(\mathbf{p})\|_\infty, \end{aligned}$$

as is stated by (25). This concludes the proof. \square

Note that Lemma 15 is true regardless of the underlying interference model $\mathcal{I}(\mathbf{p})$, e.g., $\mathcal{I}(\mathbf{p})$ does not need to be monotonic. We will use Lemma 15 in a simplified form as follows:

Corollary 16. *Suppose optimization problem (5) fulfill qualifying condition ($Q_{3.b}$) up to a scaling factor $\alpha > 0$, i.e., if*

$$\alpha \|\mathcal{I}(\mathbf{p})\|_\infty < \frac{\mathbf{q}(\mathbf{p})}{1 + \mathbf{q}(\mathbf{p})}.$$

Then, optimization problem (23) is Fast-Lipschitz if

$$\max_{i,z} |\Omega_i(z)| \leq \alpha.$$

This corollary allows us to say that problem (23), regardless the underlying interference model $\mathcal{I}(\mathbf{p})$, is Fast-Lipschitz if

$$\max_{i,z} |\Omega_i(z)| < \frac{1}{\|\mathcal{I}(\mathbf{p})\|_\infty} \frac{\mathbf{q}(\mathbf{p})}{1 + \mathbf{q}(\mathbf{p})} \quad \forall \mathbf{p}. \quad (26)$$

For fading coefficients from an arbitrary distribution, the function $\Omega_i(z)$ in equation (24) might not be expressed on closed form. However, the max-value of $\Omega_i(z)$ can be found through numerical calculations. We now apply Corollary 16 to two different distributions of the channel fading Θ , one is analyzed analytically and one is studied numerically.

A. Fading models

In what follows we consider two different fading models. First we investigate the case where the channel fading coefficient Θ follows a Rayleigh distribution, whereby the worst-case value of Ω can be determined analytically. Thereafter, we investigate the case when Θ follows an exponential distribution. In this case we find the worst-case value of Ω through numeric calculation.

1) *Rayleigh distribution:* Assume Θ_i is follows a Rayleigh distribution with parameter λ_i and with pdf

$$\theta_i(y) = \frac{y}{\lambda_i^2} e^{-y^2/2\lambda_i^2}, \quad \lambda_i > 0. \quad (27)$$

Recalling the definition of $\Omega_i(z)$ in (24), we calculate the first term of $\Omega_i(z)$ as

$$\int_{\mathbf{z}} \frac{\theta_i(y)}{y} dy = \int_{\mathbf{z}} \frac{e^{-y^2/2\lambda_i^2}}{\lambda_i^2} dy.$$

By the substitution $y = \sqrt{2}\lambda_i t$ we get $dy = \sqrt{2}\lambda_i dt$ and

$$\int_{\mathbf{z}} \frac{\theta_i(y)}{y} dy = \int_{z/\sqrt{2}\lambda_i}^{\infty} \frac{e^{-t^2}}{\lambda_i^2} \sqrt{2}\lambda_i dt = \sqrt{\frac{\pi}{2\lambda_i^2}} \text{erfc}\left(\frac{z}{\sqrt{2}\lambda_i}\right),$$

where $\text{erfc}(\cdot)$ is the complementary error function. Therefore, we have

$$\Omega_i(z) = \sqrt{\frac{\pi}{2\lambda_i^2}} \text{erfc}\left(\frac{z}{\sqrt{2}\lambda_i}\right) - \frac{z}{\lambda_i^2} e^{-z^2/2\lambda_i^2}.$$

and

$$\frac{d\Omega_i(z)}{dz} = \frac{e^{-z^2/2\lambda_i^2}}{\lambda_i^2} \left(2 - \frac{z^2}{\lambda_i^2}\right),$$

which is smooth, and equal to zero only when $z = \sqrt{2}\lambda_i$. Therefore, the extreme values of $\Omega_i(z)$ must occur as $z \rightarrow 0$, $z = \sqrt{2}\lambda_i$, or $z \rightarrow \infty$. Evaluating Ω_i at these points gives $\Omega_i(z) \rightarrow \sqrt{\pi/2} \frac{1}{\lambda_i}$ as $z \rightarrow 0$,

$$\Omega_i(\sqrt{2}\lambda_i) = \frac{1}{\lambda_i} \underbrace{\left(\sqrt{\frac{\pi}{2}} \text{erfc}(1) - \sqrt{2}e^{-1}\right)}_{\approx -0.323} > -\frac{1}{3\lambda_i},$$

and $\Omega_i(z) \rightarrow 0$ as $z \rightarrow \infty$ respectively. It follows that $\max_{i,z} |\Omega_i(z)| \leq \alpha$ if

$$\alpha \geq \max \left\{ \sqrt{\pi/2} \frac{1}{\lambda_i}, \frac{1}{3} \frac{1}{\lambda_i} \right\} \Leftrightarrow \lambda_i \geq \frac{\sqrt{\pi/2}}{\alpha}$$

for all i . This means that if

- a) the original (deterministic and outage-free) problem (5) is Fast-Lipschitz by qualifying condition \mathbf{Q}_3 , i.e., $\alpha \leq 1$ in Corollary 16, and
 - b) the channel fading Θ_i follows a Rayleigh distribution (27) with parameter $\lambda_i \geq \sqrt{\pi/2}$,
- then problem (23) is Fast-Lipschitz by Corollary 16. It follows that the iterations (22) converge to \mathbf{p}^* , and \mathbf{p}^* is the unique optimal solution of the optimization problem (23).

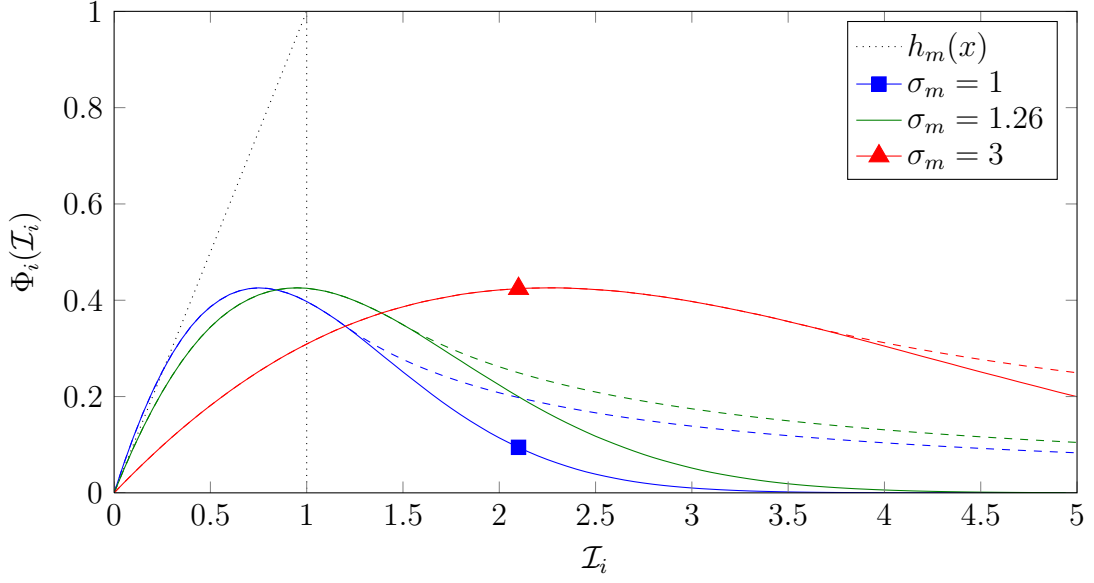


Figure 1. The *smoothed mobile behavior function* $\Phi(\mathcal{I}_i)$ for different λ_i , when Θ_i follows a Rayleigh distribution. The dashed lines show the best approximations that are absolutely subhomogeneous, as required in [35]. The dotted line shows the function $h(x)$ when $b = 1$.

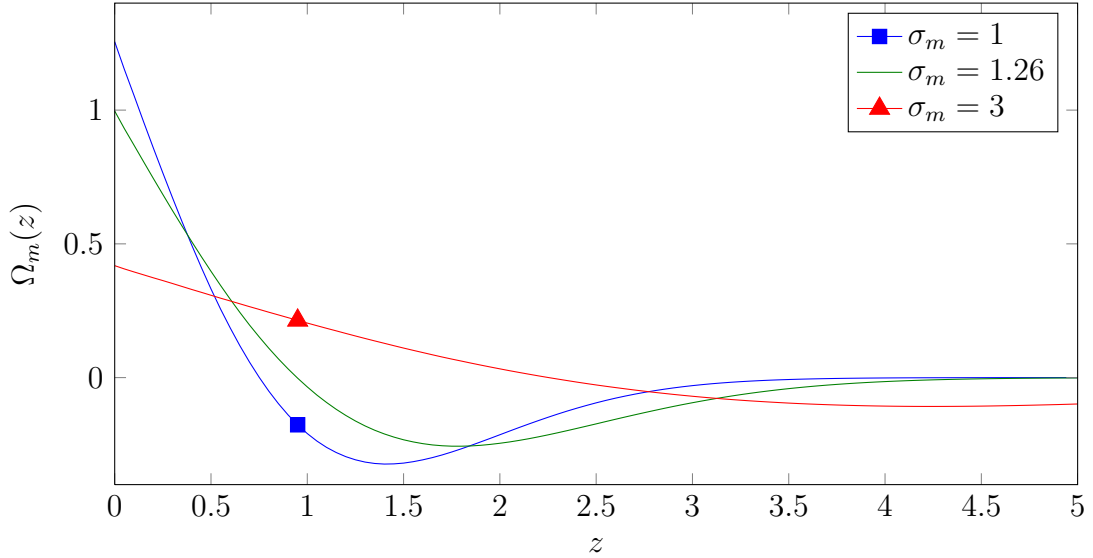


Figure 2. This figure show the behaviour of $\Omega_i(z)$ for different λ_i . When $\alpha = 1$, $\lambda_i = \sqrt{\pi/2} \approx 1.26$ is the lower limit of λ_i for which Corollary 16 applies (i.e., $|\Omega_i(z)| < 1 \forall z$).

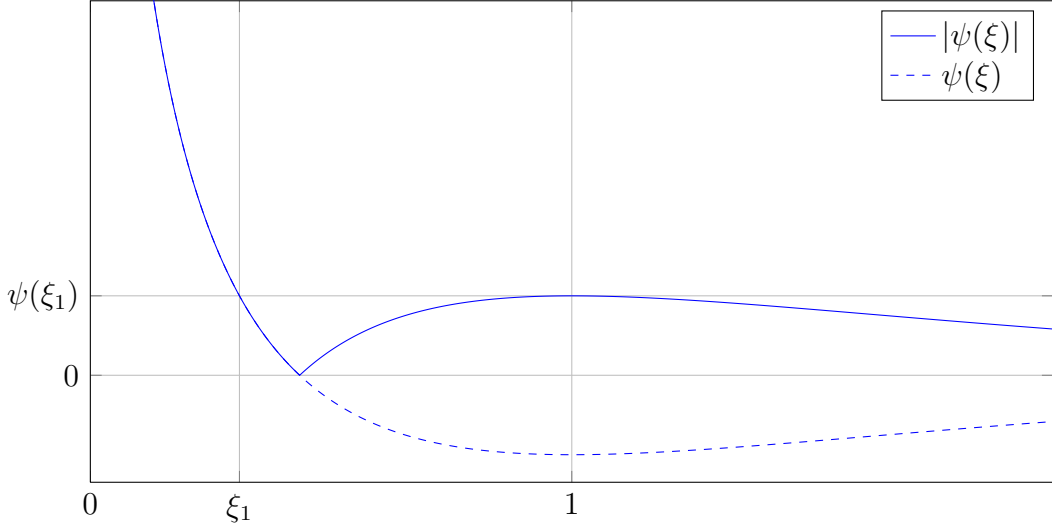


Figure 3. Graph of $\psi(\xi)$ in equation (29).

2) *Exponential distribution:* We now given an application of Corollary 16 to the case when the channel fading coefficients Θ_i are exponentially distributed,

$$\Theta_i \sim \theta_i(y | \lambda) = \lambda e^{-\lambda y}. \quad (28)$$

This is know as Rayleigh fading. Denote $z \triangleq \mathcal{I}(\mathbf{p})/b$ (we will drop the transmitter index i to get a clearer notation), and highlight the λ -dependence of Ω by writing

$$\Omega(z, \lambda) = \int_z^\infty \frac{\theta(y)}{y} dy - \theta(z | \lambda) = \lambda \left(\int_{\lambda z}^\infty \frac{e^{-t}}{t} dt - e^{-\lambda z} \right) = \lambda \psi(\lambda z)$$

where

$$\psi(\xi) \triangleq \int_\xi^\infty \frac{e^{-t}}{t} dt - e^{-\xi}. \quad (29)$$

The function $\psi(\xi)$ is shown in Figure 3.

To use the result in (26) we must show that the absolute value of Ω is small enough. We will see that this is typically the case, except when $z = \mathcal{I}/b$ goes to zero. This cannot happen in practice, since the non-zero background noise η always lower bounds the interference. Therefore, we assume that $z = \mathcal{I}/b$ is lower bounded by some z_{\min} . For any given lower bound z_{\min} , introduce

$$\sigma_{z_{\min}}(\lambda) = \max_{z \geq z_{\min}} |\Omega(z, \lambda)|.$$

The function $\sigma_{z_{\min}}$ is the worst case value over all possible values of λ , given that $z \geq z_{\min}$.

To find $\sigma_{z_{\min}}(\lambda)$, let $\xi = \lambda z$, whereby

$$|\Omega(z, \lambda)| = \lambda |\psi(\lambda z)| = \lambda |\psi(\xi)|.$$

For a fixed λ , it is sufficient to find the $z \geq z_{\min}$ that maximizes $|\psi(\lambda z)|$ or, equivalently, the $\xi \geq \xi_{\min} = \lambda z_{\min}$ that maximizes $|\psi(\xi)|$. Consider the plot of $\psi(\xi)$ is shown in Figure 3. The derivative

$$\frac{d\psi}{d\xi} = e^{-\xi} \left(1 - \frac{1}{\xi}\right)$$

is zero only when $\xi = 1$, and the second derivative is always positive. The dashed lines highlight where $\xi = 1$ and

$$\xi = \xi_1 = \{t : \psi(t) = -\psi(1)\}.$$

In order to maximize $|\psi(\xi)|$, it is clear that ξ should be chosen as

$$\xi = \begin{cases} 1 & \text{if } \xi_1 \leq \xi_{\min} \leq 1, \\ \xi_{\min} & \text{otherwise.} \end{cases}$$

In terms of the variables λ and z we therefore have

$$\sigma_{z_{\min}}(\lambda) = \max_{z \geq z_{\min}} |\Omega(z, \lambda)| = \begin{cases} \lambda\psi(\lambda z_{\min}) = \Omega(z_{\min}, \lambda), & \text{if } \lambda < \frac{\xi_1}{z_{\min}}, \\ -\lambda\psi(1), & \text{if } \frac{\xi_1}{z_{\min}} \leq \lambda \leq \frac{1}{z_{\min}}, \\ -\lambda\psi(\lambda z_{\min}) = -\Omega(z_{\min}, \lambda), & \text{if } \frac{1}{z_{\min}} < \lambda. \end{cases}$$

It is clear that any stationary point of $\sigma_{z_{\min}}$ must also be a stationary point of $\Omega(z_{\min}, \lambda)$, with derivative

$$\begin{aligned} \frac{d\Omega(z_{\min}, \lambda)}{d\lambda} &= \frac{d}{d\lambda} (\lambda\psi(\lambda z_{\min})) = \psi(\lambda z_{\min}) + \lambda \frac{d\psi(\lambda z_{\min})}{d\lambda} \\ &= \psi(\lambda z_{\min}) + \lambda \left(z_{\min} e^{-\lambda z_{\min}} \left(1 - \frac{1}{\lambda z_{\min}}\right) \right) \\ &= \psi(\lambda z_{\min}) + e^{-\lambda z_{\min}} (\lambda z_{\min} - 1). \end{aligned}$$

Setting the expression above to zero and solving numerically gives the two solutions

$$\begin{cases} \lambda z_{\min} = v_1 \approx 0.1184 & \text{and} \\ \lambda z_{\min} = v_2 \approx 1.5656, \end{cases}$$

i.e., when $\lambda = v_1/z_{\min}$ and $\lambda = v_2/z_{\min}$. Inserting these values into $\sigma_{z_{\min}}(\lambda)$ gives the values

$$\begin{cases} \sigma_{z_{\min}}\left(\frac{v_1}{z_{\min}}\right) = \frac{v_1}{z_{\min}} \psi(v_1) \approx \frac{0.093}{z_{\min}} & \text{and} \\ \sigma_{z_{\min}}\left(\frac{v_2}{z_{\min}}\right) = \frac{v_2}{z_{\min}} \psi(v_2) \approx \frac{0.185}{z_{\min}} \end{cases}$$

of the two local maxima shown in the Figure 4. Assuming $z \geq z_{\min}$, we therefore have

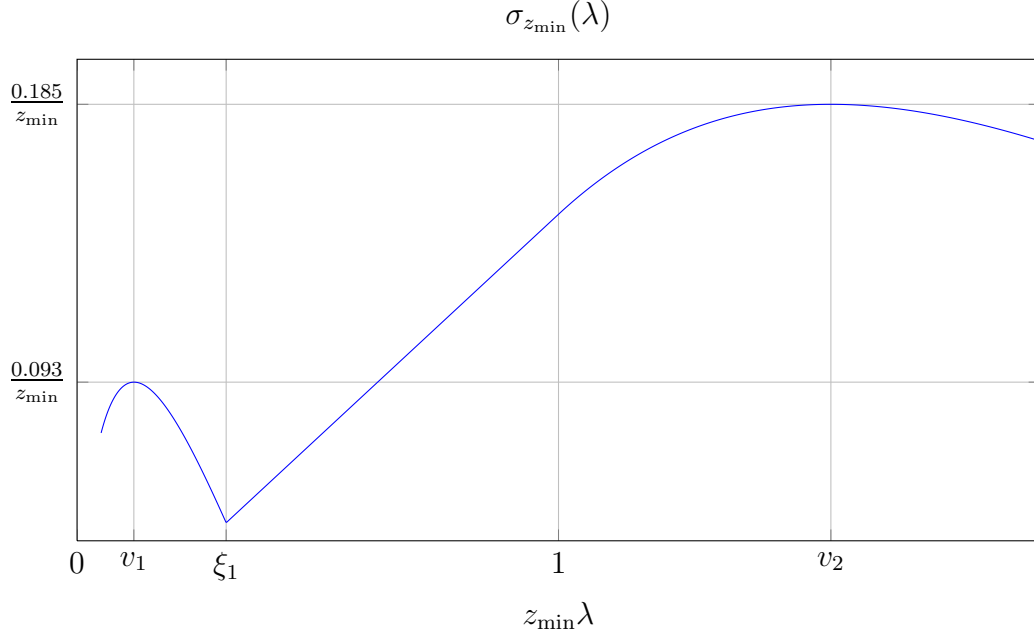


Figure 4. Plot of $\sigma_{z_{\min}}(\lambda)$, note that the x-axis is scaled by z_{\min} .

$$|\Omega(z, \lambda)| \leq \max_{z \geq z_{\min}} |\Omega(\lambda, z)| = \sigma_{z_{\min}}(\lambda) \leq 0.185/z_{\min}$$

for any parameter value λ of the fading coefficient distribution parameter. In particular, Corollary 16 states that problem (23) is Fast-Lipschitz if

$$\|\nabla \mathcal{I}(\mathbf{p})\|_{\infty} < \frac{0.185}{z_{\min}} \frac{\mathbf{q}(\mathbf{p})}{1 + \mathbf{q}(\mathbf{p})}$$

for all $\mathbf{p} \geq 0$, where

$$\mathbf{q}(\mathbf{p}) = \min_j \frac{\min_i \nabla_i \kappa_j(\mathbf{p})}{\max_i \nabla_i \kappa_j(\mathbf{p})}$$

is given by the characteristics of the cost function $\kappa(\mathbf{p})$.

This example has showed how problems without monotonicity properties can be analyzed with Fast-Lipschitz optimization. The price one has to pay to ensure optimality is the tighter bound on $\|\nabla \mathcal{I}(\mathbf{p})\|_{\infty}$ (note that $\mathbf{q}(\mathbf{p})/(1 + \mathbf{q}(\mathbf{p})) \leq 1/2$), as opposed to requiring $\|\nabla \mathcal{I}(\mathbf{p})\|_{\infty} < 1$ for monotonic functions, which is sufficient to show contractivity.

VI. CONCLUSIONS AND FUTURE WORK

In this paper we examined the conditions under which power control algorithms with standard, type-II standard and more general functions fall under the Fast-Lipschitz framework. This allowed us to give the studied problems a richer notion of optimality. In the process

we established a new qualifying condition for Fast-Lipschitz optimization that shows a close relation between requirements on the cost function and requirements on the constraints to achieve optimality.

In this paper we assumed that the functions are everywhere differentiable. This is not necessarily required by the standard and type-II standard formulations, and we believe this requirement can be dropped also in Fast-Lipschitz optimization. However, this is something that still needs to be formalized. Furthermore, the results of Section IV.C hint of possible relaxations of the qualifying conditions if one considers cones different from the non-negative orthant.

APPENDIX: FAST-LIPSCHITZ QUALIFYING CONDITIONS

Given a problem on Fast-Lipschitz form (7), the General Qualifying Condition (GQC) of Table 1 guarantees that the problem is Fast-Lipschitz.

Theorem 17 ([? , Theorem 7]). *Assume problem (7) is feasible, and that the General Qualifying Conditions GQC in Table 1 hold for every $\mathbf{x} \in \mathcal{D}$. Then, the problem is Fast-Lipschitz, i.e.,*

$$\mathbf{x}^{k+1} := \mathbf{f}(\mathbf{x}^k)$$

converges to $\mathbf{x}^ = \mathbf{f}(\mathbf{x}^*)$, and \mathbf{x}^* is the unique Pareto optimal solution of problem (7).*

There are several special cases of GQC that more convenient easier to work with. We list three of them (the ones used in this paper) in Table 1.

Proposition 18 ([?]). *If any of qualifying conditions Q_1 - Q_3 hold, then so does GQC.*

Remark 19. Note that the qualifying conditions only are sufficient, not necessary. This means that there can be problems that are Fast-Lipschitz but fail to fulfill the qualifying conditions of Table 1.

REFERENCES

- [1] R. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 1341–1347, Sept. 1995.

General Qualifying Condition	
GQC	<p>(GQC.a) $\nabla f_0(\mathbf{x}) \geq \mathbf{0}$ with non-zero rows</p> <p>(GQC.b) $\ \nabla f(\mathbf{x})\ < 1$</p> <p>There exists a $k \in \{1, 2, \dots\} \cup \infty$ such that</p> <p>(GQC.c) when $k < \infty$, then $\nabla f(\mathbf{x})^k \geq \mathbf{0}$</p> <p>(GQC.d) when $k > 1$, then $\left\ \sum_{l=1}^{k-1} \nabla f(\mathbf{x})^l \right\ _\infty < \mathbf{q}(\mathbf{x}) \triangleq \min_j \frac{\min_i [\nabla f_0(\mathbf{x})]_{ij}}{\max_i [\nabla f_0(\mathbf{x})]_{ij}}$</p>
Qualifying Condition 1	
Q₁	<p>(Q₁.a) $\nabla f_0(\mathbf{x}) \geq \mathbf{0}$ with non-zero rows</p> <p>(Q₁.b) $\ \nabla f(\mathbf{x})\ < 1$</p> <p>(Q₁.c) $\nabla f(\mathbf{x}) \geq \mathbf{0}$</p>
Qualifying Condition 2	
Q₂	<p>(Q₂.a) $\nabla f_0(\mathbf{x}) > \mathbf{0}$</p> <p>(Q₂.b) $\nabla f(\mathbf{x})^2 \geq \mathbf{0}$, (e.g., $\nabla f(\mathbf{x}) \leq \mathbf{0}$)</p> <p>(Q₂.c) $\ \nabla f(\mathbf{x})\ _\infty < \mathbf{q}(\mathbf{x}) \triangleq \min_j \frac{\min_i [\nabla f_0(\mathbf{x})]_{ij}}{\max_i [\nabla f_0(\mathbf{x})]_{ij}}$</p>
Qualifying Condition 3	
Q₃	<p>(Q₃.a) $\nabla f_0(\mathbf{x}) > \mathbf{0}$</p> <p>(Q₃.b) $\ \nabla f(\mathbf{x})\ _\infty < \frac{\mathbf{q}(\mathbf{x})}{1 + \mathbf{q}(\mathbf{x})}$</p>

Table 1

FAST-LIPSCHITZ QUALIFYING CONDITIONS FROM [?] . QUALIFYING CONDITIONS 1-3 IMPLY THE GENERAL CONDITION GQC.

- [2] F. Meshkati, H. Poor, and S. Schwartz, “Energy-efficient resource allocation in wireless networks,” *IEEE Signal Process. Mag.*, vol. 24, pp. 58–68, May 2007.
- [3] J. Zander, “Distributed cochannel interference control in cellular radio systems,” *IEEE Trans. Veh. Technol.*, vol. 41, pp. 305–311, Aug. 1992.
- [4] S. Grandhi, R. Vijayan, and D. Goodman, “Distributed power control in cellular radio systems,” *IEEE Trans. Commun.*, vol. 42, pp. 226–228, Feb. 1994.

- [5] G. Foschini and Z. Miljanic, “A simple distributed autonomous power control algorithm and its convergence,” *IEEE Trans. Veh. Technol.*, vol. 42, pp. 641–646, Nov. 1993.
- [6] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.
- [7] D. Mitra, “An asynchronous distributed algorithm for power control in cellular radio systems,” in *Proc. of 4th WINLAB Workshop*, 1993.
- [8] K.-K. Leung, C.-W. Sung, W. S. Wong, and T. Lok, “Convergence theorem for a general class of power-control algorithms,” *IEEE Trans. Commun.*, vol. 52, pp. 1566–1574, Sept. 2004.
- [9] C.-W. Sung and W. S. Wong, “A distributed fixed-step power control algorithm with quantization and active link quality protection,” *IEEE Trans. Veh. Technol.*, vol. 48, pp. 553–562, Mar. 1999.
- [10] C. Sung and K. Leung, “A generalized framework for distributed power control in wireless networks,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 2625–2635, July 2005.
- [11] S. Ulukus and R. Yates, “Stochastic power control for cellular radio systems,” *IEEE Trans. Commun.*, vol. 46, pp. 784–798, June 1998.
- [12] J. Luo, S. Ulukus, and A. Ephremides, “Standard and quasi-standard stochastic power control algorithms,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 2612–2624, July 2005.
- [13] M. Biguesh and S. Gazor, “Distributed power control in cellular communication systems concerning inaccurate SINR reports,” *IEEE Trans. Veh. Technol.*, vol. 60, pp. 3657–3666, Oct. 2011.
- [14] C. Fischione, “Fast-Lipschitz optimization with wireless sensor networks applications,” *IEEE Trans. Autom. Control*, vol. 56, pp. 2319–2331, Oct. 2011.
- [15] J. Tsitsiklis, “Problems in decentralized decision making and computation,” Ph.D. dissertation, 1984.
- [16] D. Palomar and M. Chiang, “Alternative distributed algorithms for network utility maximization: Framework and applications,” *IEEE Trans. Autom. Control*, vol. 52, pp. 2254–2269, Dec. 2007.
- [17] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. and Trends® in Mach. Learning*, vol. 3, no. 1, pp. 1–122, July 2011.
- [18] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE Trans. Inf. Theory*, vol. 52, pp. 2508–2530, June 2006.

- [19] A. Olshevsky and J. Tsitsiklis, “Convergence speed in distributed consensus and averaging,” *SIAM J. Control and Optimization*, vol. 48, no. 1, pp. 33–55, Feb. 2009.
- [20] A. Nedić, A. Ozdaglar, and P. Parrilo, “Constrained consensus and optimization in multi-agent networks,” *IEEE Trans. Autom. Control*, vol. 55, pp. 922–938, Apr. 2010.
- [21] H. Boche and M. Schubert, “The structure of general interference functions and applications,” *IEEE Trans. Inf. Theory*, vol. 54, pp. 4980–4990, Nov. 2008.
- [22] H. Feyzmahdavian, M. Johansson, and T. Charalambous, “Contractive interference functions and rates of convergence of distributed power control laws,” *IEEE Trans. Wireless Commun.*, vol. 11, pp. 4494–4502, Dec. 2012.
- [23] H. Ji and C.-Y. Huang, “Non-cooperative uplink power control in cellular radio systems,” *Wireless Networks*, vol. 4, no. 3, pp. 233–240, July 1998.
- [24] C. Saraydar, N. Mandayam, and D. Goodman, “Efficient power control via pricing in wireless data networks,” *IEEE Trans. Commun.*, vol. 50, Feb. 2002.
- [25] M. Xiao, N. Shroff, and E. K. P. Chong, “A utility-based power-control scheme in wireless cellular systems,” *IEEE/ACM Trans. Netw.*, vol. 11, pp. 210–221, Apr. 2003.
- [26] F. Li and L. Wang, “A dynamic game algorithm for power allocation in cognitive relay transmission,” *IEEE Trans. Commun.*, vol. 61, pp. 4440–4447, Nov. 2013.
- [27] H. Tuy, “Monotonic optimization: Problems and solution approaches,” *SIAM J. Optimization*, vol. 11, no. 2, pp. 464–494, 2000.
- [28] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, “Robust monotonic optimization framework for multicell miso systems,” *IEEE Trans. Signal Process.*, vol. 60, pp. 2508–2523, May 2012.
- [29] Y. J. A. Zhang, L. Qian, and J. Huang, “Monotonic optimization in communication and networking systems,” *Found. and Trends® in Networking*, vol. 7, no. 1, pp. 1–75, 2013.
- [30] G. Notarstefano and F. Bullo, “Distributed abstract optimization via constraints consensus: Theory and applications,” *IEEE Trans. Autom. Control*, vol. 56, pp. 2247–2261, Oct. 2011.
- [31] M. Bürger, G. Notarstefano, F. Allgöwer, and F. Bullo, “A distributed simplex algorithm for degenerate linear programs and multi-agent assignments,” *Automatica*, vol. 48, no. 9, pp. 2298–2304, Sept. 2012.
- [32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [33] A. Möller and U. Jönsson, “Input-output analysis of power control in wireless networks,” *IEEE Trans. Autom. Control*, vol. 58, pp. 834–846, Apr. 2013.

- [34] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, “A tutorial on geometric programming,” *Optimization and Engineering*, vol. 8, no. 1, pp. 67–127, Mar. 2007.
- [35] C. Nuzman, “Contraction approach to power control, with non-monotonic applications,” in *IEEE Global Telecommun. Conf., (GLOBECOM)*, Nov. 2007.